

Review of the Master Thesis

Author: Dinh Le Thanh

Title: Question and Answer Classifier for Closed Domain Interactive Question Answering

Supervisor: RNDr. Markéta Lopatková, Ph. D.

Description

The thesis tackles the issues of the classification of question types needed in Question Answering systems, and the classification of follow up questions into topic continuation and topic shift in Interactive Question Answering.

There are 7 Chapters in the thesis: Chapter I provides a general insight into studied problems. Research issues and objectives are clearly described and well defined.

Chapter II depicts classification problem within the field of Question Answering, gives short yet well arranged introduction and history with mentioning related works. The chapter also explains the taxonomy and enumerates advantages and disadvantages of two different approaches in classification, namely the rule-based approach and the machine learning one.

Chapter III introduces the “topic shift“ and the “topic continuation“ task. Once again, the presented introduction is eligibly readable and it is accompanied with previous work on the task.

Chapter IV describes two classifiers used in the work: the Naïve Bayes classifier (for question classificaton) and Decision tree (for identification of topic and follow-up quesions). Evaluation measures (Precision, Recall, F-measure) are also explained there.

The most important part of the work can be found in Chapters V and VI: Chapter V depicts experiments done for question classification. Both LR-TREC corpus together with taxonomy established by Li and Roth (2002) and BOB-QC corpus with new taxonomy prepared by the author are described. For training purposes, LR-TREC data are always used. For testing purposes and for obtaining tables of results, both pieces of data (separately) are used. Number of experiments is doubled by two options in the setting taxonomy (a fine-grained or coarse classification). All experiments are well documented, especially concerning feature selection, and afterwards analysed. Finally, a short comparison of results obtained on the different corpora is attached.

Chapter VI is structured similarly to the previous Chapter and it is dedicated to the continuation and topic shift detection.

Chapter VII clearly summarises the presented work. It nicely brings conclusions from experiments and suggests some possible steps of futher work that can profit from the thesis.

On the whole, the text is written in good English. It is clear and can be used as an introduction into the Question Answering classification. It shows itself that the author not only understands the topic, but on the top of that he is capable to divide work into individual steps and describe all important parts very thoroughly. The amount of work behind the presented text is tremendous. The author had to understand and to work with several third party packages and tools (i.e. package for similarity, external parser and tagger).

Minor drawbacks and questions

- I do not understand the difference mentioned in properties of rule-based approach („maximum flexibility“) on p. 15 vs. machine learning approach („more flexible“) on p. 16.
- I am a bit sceptical about the comparison of results tested on different data on p. 38. Can we really compare the obtained evaluation values? How can we interpret the difference?
- There are 4 similarities based on noun on p. 43, two based on noun and two based on verb are supposed to be there.
- Based on all Tables 12-15 on p. 44, I suppose that either the numbers or the label of Table 16 are inconsistent with the previous text. How can we read and interpret the results, mainly while comparing with related works?
- During training I miss a mention about the tuning of parameters of classifiers, mainly for Decision tree. Was the tuning done? If so, can you measure the impact of the tuning?

Conclusion

The amount and the importance of my objections are very tenuous. All that has been mentioned above proves that the author can solve NLP tasks independently. In my honest opinion, the work complies with the requirements for a Master Thesis at MFF – **I recommend to accept the thesis for the defence.**

Prague, September 7, 2009

Mgr. Pavel Schlesinger
Institute of Formal and Applied Linguistics
Charles University in Prague

Review of the Master Thesis

Author: Dinh Le Thanh

Title: Question and Answer Classifier for Closed Domain Interactive Question Answering

Supervisor: RNDr. Markéta Lopatková, Ph. D.

Description

The thesis tackles the issues of the classification of question types needed in Question Answering systems, and the classification of follow up questions into topic continuation and topic shift in Interactive Question Answering.

There are 7 Chapters in the thesis: Chapter I provides a general insight into studied problems. Research issues and objectives are clearly described and well defined.

Chapter II depicts classification problem within the field of Question Answering, gives short yet well arranged introduction and history with mentioning related works. The chapter also explains the taxonomy and enumerates advantages and disadvantages of two different approaches in classification, namely the rule-based approach and the machine learning one.

Chapter III introduces the “topic shift” and the “topic continuation” task. Once again, the presented introduction is eligibly readable and it is accompanied with previous work on the task.

Chapter IV describes two classifiers used in the work: the Naïve Bayes classifier (for question classificaton) and Decision tree (for identification of topic and follow-up questions). Evaluation measures (Precision, Recall, F-measure) are also explained there.

The most important part of the work can be found in Chapters V and VI: Chapter V depicts experiments done for question classification. Both LR-TREC corpus together with taxonomy established by Li and Roth (2002) and BOB-QC corpus with new taxonomy prepared by the author are described. For training purposes, LR-TREC data are always used. For testing purposes and for obtaining tables of results, both pieces of data (separately) are used. Number of experiments is doubled by two options in the setting taxonomy (a fine-grained or coarse classification). All experiments are well documented, especially concerning feature selection, and afterwards analysed. Finally, a short comparison of results obtained on the different corpora is attached.

Chapter VI is structured similarly to the previous Chapter and it is dedicated to the continuation and topic shift detection.

Chapter VII clearly summarises the presented work. It nicely brings conclusions from experiments and suggests some possible steps of futher work that can profit from the thesis.

On the whole, the text is written in good English. It is clear and can be used as an introduction into the Question Answering classification. It shows itself that the author not only understands the topic, but on the top of that he is capable to divide work into individual steps and describe all important parts very thoroughly. The amount of work behind the presented text is tremendous. The author had to understand and to work with several third party packages and tools (i.e. package for similarity, external parser and tagger).

Minor drawbacks and questions

- I do not understand the difference mentioned in properties of rule-based approach („maximum flexibility“) on p. 15 vs. machine learning approach („more flexible“) on p. 16.
- I am a bit sceptical about the comparison of results tested on different data on p. 38. Can we really compare the obtained evaluation values? How can we interpret the difference?
- There are 4 similarities based on noun on p. 43, two based on noun and two based on verb are supposed to be there.
- Based on all Tables 12-15 on p. 44, I suppose that either the numbers or the label of Table 16 are inconsistent with the previous text. How can we read and interpret the results, mainly while comparing with related works?
- During training I miss a mention about the tuning of parameters of classifiers, mainly for Decision tree. Was the tuning done? If so, can you measure the impact of the tuning?

Conclusion

The amount and the importance of my objections are very tenuous. All that has been mentioned above proves that the author can solve NLP tasks independently. In my honest opinion, the work complies with the requirements for a Master Thesis at MFF – **I recommend to accept the thesis for the defence.**

Prague, September 7, 2009

Mgr. Pavel Schlesinger
Institute of Formal and Applied Linguistics
Charles University in Prague